

GLAZE: Protecting Artists from Style Mimicry by Text-to-Image Models

Year: 2023 | Citations: 235 | Authors: Shawn Shan, Jenna Cryan, Emily Wenger, Haitao Zheng, Rana Hanocka

Abstract

Recent text-to-image diffusion models such as MidJourney and Stable Diffusion threaten to displace many in the professional artist community. In particular, models can learn to mimic the artistic style of specific artists after "fine-tuning" on samples of their art. In this paper, we describe the design, implementation and evaluation of Glaze, a tool that enables artists to apply "style cloaks" to their art before sharing online. These cloaks apply barely perceptible perturbations to images, and when used as training data, mislead generative models that try to mimic a specific artist. In coordination with the professional artist community, we deploy user studies to more than 1000 artists, assessing their views of AI art, as well as the efficacy of our tool, its usability and tolerability of perturbations, and robustness across different scenarios and against adaptive countermeasures. Both surveyed artists and empirical CLIP-based scores show that even at low perturbation levels ($p=0.05$), Glaze is highly successful at disrupting mimicry under normal conditions (>92%) and against adaptive countermeasures (>85%).